#8

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): Egger, et al. Attorney Docket No.: 4256

Serial No.: 08/649,304 Examiner: Amsbury, W.

Filed: May 17, 1996 Group Art Unit: 2307

For: METHOD AND APPARATUS FOR INDEXING, SEARCHING AND DISPLAYING DATA. This application is a continuation-in-part of US Application No. 08/076,658, filed June 14, 1993.

---

enter 1/27/98

INFORMATION DISCLOSURE STATEMENT

The Assistant Commissioner for Patents
Washington, DC 20231

Dear Sir:

Pursuant to sections 1.56, 1.97 and 1.98 of Title 37 of the Code of Federal Regulations, the following information and comments are set forth in compliance with the applicant's duty of disclosure, as well as the duty of others substantively involved in the preparation of the above-titled application:

I. Patents

Applicants, and others substantively involved in the preparation of the application, are aware of the following patents, abstracts of which are enclosed:

a) US Patent No. 5,265,065, issued to Howard R. Turtle on November 23, 1993, and assigned to West Publishing Company, entitled "Method and Apparatus

for Information Retrieval from a Database by Replacing Domain Stemmed Phrases in a Natural Language to Create a Query Search."

II. Publications

Applicants, and others substantively involved in the preparation of the application, are aware of the following publications, copies of which are enclosed:

a) Agosti, Colotti, Gradenigo, A Two-Level Hypertext Retrieval Model for Legal Data, SIGIR '91 (1991).

b) Fowler, Fowler, Wilson, Integrating Query, Thesaurus and Documents Through a Common Visual Representation, SIGIR '91 (1991).

c) Rose, Daniel E., & Belew, Richard K., Legal Information Retrieval: a hybrid approach, ICAIL '89 (1989).

d) Belew, Richard, A Connectionist Approach to Conceptual Information Retrieval, ICAIL '87 (1987).

e) Gelbart, D & Smith, J.C., Beyond Boolean Search: FLEXICON, A Legal Text-Based Intelligent System, ICAIL '91 (1991).

f) Lin, A Self-Organizing Semantic Map for Information Retrieval, SIGIR '91 (1991).

g) Turtle, Howard R., & Croft, W. Bruce, Inference Networks for Document Retrieval, SIGR '90 (1990).

III. Other

Applicants, and others substantively involved in the preparation of the application, are aware of the following additional information:

a) ConQuest Corporation is marketing a full-text retrieval system. However, Applicant is not in possession of any published documentation on the ConQuest system.

IV. Background of the References

The references located primarily relate to use of text-by-text analysis procedures (boolean searches) to scan and retrieve items from a database. Text-by-text analysis procedures are disclosed in the background section of the application. Many experts in the field of full text information retrieval refer to text-by-text analysis procedures as procedures which operate under the "semantic model". The semantic model is based on the premise that documents with words in common are likely to be related. Most (if not all) full text information retrieval systems operate under this premise and therefore use the semantic model approach.

There are variations on the specific search techniques that can be used under the semantic model approach. For example, neural net systems also known as "connectionist" systems (such as publications e, c.1 and c.2 above) use one technique for text-by-text analysis. In contrast, the novel search techniques claimed in the above-titled application do not rely on the semantic model approach and are not described in any of the references.

V. A Two-Level Hypertext Retrieval Model for Legal Data

This reference discloses a semantic model approach to text searching. The article describes searching for the desired word or words and related words using a "thesaurus" concept. The developers of this system created and entered into a database a thesaurus of related terms for text searching. Not only can a text search of an exact word be conducted, but predetermined synonyms of the desired word may also be searched. The method disclosed in this reference is relevant because it discloses the use of hypertext links (based on Hypercard™). However, there is no analysis of the links nor any attempt to derive any knowledge of the textual database through analysis of the links. In addition, there is no attempt to use visualization techniques to represent information in the database.

VI. Integrating Query, Thesaurus and Documents Through a Common Visual Representation

The system disclosed in this reference is relevant because it uses an iterative process to refine searches and the system uses graphics. However, the system provides a graphical method to do what other semantic modeled systems do: refine the search query by adding and deleting textual terms. The non-semantic search techniques of the claimed invention are different than the semantic search techniques disclosed in the article. Also, both the information being displayed, semantic information (related words), and the visual representation are fundamentally different from the claimed invention. The claimed invention offers

graphical representations of information about relationships between documents and portions of documents not relationships between search words.

VII. Legal Information Retrieval: a Hybrid Approach and A Connectionist Approach to Conceptual Information Retrieval

The hybrid legal information retrieval system described in the 1989 article is relevant because it is a different approach than other semantic models. The hybrid approach, like many others, imposes a higher-level semantic structure on a text database. Unlike others, the semantic structure described combines a standard symbolic representation (named nodes and typed links) with a connectionist model (or neural network). The goal of this system is to learn new database relationships as an artifact of using the system.

As the hybrid system is used, it learns database relations through feedback. However, there is no description of deriving higher-order knowledge of the text through analysis of the inherent text links.

The system is also relevant because it is graphics oriented. However, like other systems, the system focuses on graphical display of the semantic structure of the knowledge base. This is very different from the use of graphic visualization techniques in the claimed system.

The 1987 Belew article (c.2) appears to document earlier work leading up to the 1989 Belew & Rose article (c.1). Therefore, the 1987 Belew article is relevant but cumulative.

VIII. Beyond Boolean Search: FLEXICON, A Legal Text-Based Intelligent System

The system described in this article, Flexicon, is relevant because the techniques described are different than the normal techniques used with the semantic model. The Flexicon system described is explicitly tailored for legal research by assigning each document a profile that is derived based on specific attributes of the legal system. For example, in creating the profile, key legal concepts are extracted from the text by matching commonly occurring legal terms and phrases which represent the desired legal concepts.

In addition, each case or statute cited in a document becomes a key phrase in the document's profile. Thus, two document profiles may be similar by virtue of the fact that each profile cites the same legal case.

However, there is no higher-order analysis of the citation links, no attempt to derive a line of related legal cases, and no attempt at clustering the database based on the links.

In addition, the only use of graphical information in the system is generation of a histogram representing the weighted score of the retrieved documents. See page 231.

IX. A Self-Organizing Semantic Map for Information Retrieval

This system described is relevant because it organizes the database without feedback from the user (through a learning algorithm). The article appears to describe a "connectionist" approach to modeling the data base using an

unsupervised neural network learning algorithm. Again, text-by-text analysis is utilized.

The article is also relevant for its description of a map interface. The system described can represent documents graphically by assigning each document to a point on a two-dimensional grid. The grid shown is divided into "neighborhoods" representing semantic concepts. Each document is assigned to one of a discrete set of points within a "neighborhood." The map interface is both conceptually and visually different than the claimed invention.

X. Interference Networks for Document Retrieval

This article is relevant because it mentions clustering and the use of another method of information retrieval, a Bayesian inference network. This article is also based on the semantic model. The inference network uses text analysis. The article does not teach high-order analysis of links. The clustering described is conducted manually.

XI. CONCLUSION

For the foregoing reasons, it is submitted that the documents disclosed with this statement should not prevent the grant of a Patent to Applicant for their invention. It is respectfully submitted that the claims are allowable.

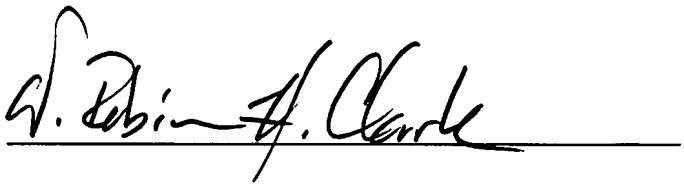
A copy of form PTO-1449 listing the above publications is attached.

The present statement reflects the direct knowledge of the undersigned and information disclosed to the undersigned by Applicant and others substantively involved in the preparation of the present application pursuant to their duty under section 1.56 of Title 37 of the Code of Federal Regulations.

As no action on the merits as yet been made, no fee is due. However, the Commissioner is hereby authorized to charge any underpayment resulting from the filing of this disclosure statement to Deposit Account No. 04-1425. A duplicate copy of this document is submitted for that purpose.

Respectfully submitted,

Dated: January 27, 1998

By: 

W. Robinson H. Clark
P-41,530

Aldo Noto
Reg. No. 35,628

Dorsey & Whitney
1330 Connecticut Avenue, N.W.
Suite 200
Washington, D.C. 20036
(202) 857-0700